HAL open science

Une expérimentation de plateforme de diffusion automatisée et collaborative des veilles avec le logiciel libre WordPress

Mathieu Andro, Mickaël Réault

To cite this version:

Mathieu Andro, Mickaël Réault. Une expérimentation de plateforme de diffusion automatisée et collaborative des veilles avec le logiciel libre WordPress. 2022. hal-03895971

HAL Id: hal-03895971

https://hal.science/hal-03895971v1

Preprint submitted on 13 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Une expérimentation de plateforme de diffusion automatisée et collaborative des veilles avec le logiciel libre WordPress

Mathieu Andro, Animateur du réseau de veille des Services du Premier ministre
Mickaël Réault, CEO Sindup

Résumé 1

Introduction 2

Méthode et description de l'expérimentation 3

Analyses 5

Une veille collaborative en mode pull 5

Vers un réseau d'organisations apprenantes 7

Un projet développé de manière agile sous une forme originale de collaboration public / privé 8

Analyses des usages 9

Analyses textuelles des contenus avec CorTexT 12

Discussion 13

Conclusion 14

Références 15

Résumé

Dans les organisations, les veilles sont encore souvent diffusées en silos informationnels via des newsletters adressées par courriels. A la faveur du confinement sanitaire et afin d'initier une veille plus partagée, ouverte et collaborative, une plateforme de diffusion automatisée des veilles depuis divers outils de surveillance comme InoReader, KB Crawl et Sindup a été expérimentée de manière agile avec le logiciel libre WordPress. Elle est accessible sur smartphone, elle propose des fonctionnalités collaboratives, sémantiques, de recherche et de capitalisation des contenus, et des statistiques détaillées de consultation. Elle a été expérimentée par une vingtaine de veilleurs d'une dizaine de services différents. La

plateforme a ensuite été hébergée, consolidée, mise en forme et sécurisée avec l'aide de l'éditeur de technologies de veille Sindup.
Ce modèle de plateforme, construit avec WordPress, est entièrement reproductible pour un très faible coût. Il pourrait permettre aux veilleurs de rendre plus efficiente la diffusion de leurs veilles et aux éditeurs de plateformes de proposer des offres moins onéreuses, plus centrées sur leur cœur de métier qui reste la mise en surveillance des sources. Il pourrait aussi permettre aux éditeurs de s'adresser à un marché plus large de PME, de TPE et d'associations.

Introduction

Au sein de nos services, la fonction veille a pour objectif de permettre aux organisations de rester à l'état de l'art dans leurs domaines, mais aussi de détecter des opportunités ou des menaces éventuelles pour leurs activités ou encore de mieux se mettre à l'écoute de la société civile. La veille documentaire s'inscrit dans le cadre stratégique de l'intelligence économique, tout comme la protection du capital informationnel et l'influence (Moreau, 2008). Les veilles sont généralement envoyées par courriels, en silos informationnels trop étanches. Elles sont encore insuffisamment numériques, transparentes, collaboratives et mutualisées.
Afin d'y remédier, un réseau de veille a été développé en 2018 au sein des Services du Premier ministre. Il s'est également progressivement élargi à l'interministériel. Ce réseau a pour objectif de favoriser les échanges de bonnes pratiques, le partage d'expérience, la capitalisation de la connaissance et la valorisation des expertises de chacun. Il est alimenté par une veille métier, des rendez-vous hebdomadaires et des réunions thématiques. Il s'intéresse aux technologies de veille, flux RSS ou web crawling, aux technologies d'analyses textuelles grâce à l'Intelligence Artificielle, au text mining, à la datavisualisation. Plus largement, il s'intéresse aussi à l'intelligence économique, à la transformation numérique, à l'open access, à l'open data, à la cyberdéfense et aux éditeurs de ressources documentaires. Il propose également des études et des analyses avec, par exemple, tout récemment, un panorama complet des outils de veille[1]. Ce réseau de veille offre aussi bien d'autres services. Il propose ainsi, en tout premier lieu, d'automatiser la surveillance des sources qui intéressent les membres du réseau via l'agrégateur de flux RSS InoReader pour lequel des formations et un accompagnement sont organisés. Mais les flux RSS ne permettent souvent de surveiller qu'une infime partie des contenus des éditeurs. Par ailleurs, certaines sources ne sont pas ou plus compatibles avec la technologie RSS. Et l'accès de nombreux contenus est contrôlé. Le réseau de veille propose de les surveiller néanmoins grâce aux robots de KB Crawl qui visitent le web à la place des veilleurs et viennent leur rapporter automatiquement les contenus qui les intéressent.
Pour finir, comme il devient humainement impossible de lire l'ensemble de la littérature produite sur certaines thématiques et difficile d'appréhender de gros corpus de textes avec objectivité, le réseau propose d'extraire des corpus sur certains sujets et de les analyser grâce aux technologies d'intelligence artificielle, de text mining et de datavisualisation avec

[1] https://plateformes-de-veille.org

CorTexT, une solution logicielle gratuite développée par la recherche scientifique française. Cette technologie permet concrètement d'identifier automatiquement les noms d'organismes, de personnes, de lieux mais aussi les concepts présents dans les textes et d'en réaliser la cartographie selon leurs occurrences et leurs cooccurrences, d'en connaître la dynamique dans le temps, mais aussi de détecter automatiquement le sentiment des textes ou des tweets…

A la faveur de la crise sanitaire, une plateforme de diffusion des veilles a été expérimentée grâce au CMS libre WordPress, dans le cadre de ce réseau de veille, puis progressivement élargie à une vingtaine de veilleurs de diverses institutions, qui souhaitaient également diffuser leurs veilles via ce média, en complément ou à la place de leurs newsletters. L'expérimentation a fait l'objet d'une publication dans la littérature professionnelle (Andro, 2021). Elle a ultérieurement été reproduite au Ministère de la Culture puis par la société Sindup qui souhaitait étudier une possible nouvelle offre de service à ses clients. A cette occasion, une collaboration avec cet éditeur de technologies de veille a été trouvée afin de reprendre l'hébergement, la sécurisation et la mise en forme de la plateforme.

Suite à un précédent article à vocation concrète et professionnelle, le présent article résulte d'une analyse plus conceptuelle et à vocation plus académique de cette expérimentation originale.

Méthode et description de l'expérimentation

Pour une description détaillée du dispositif, nous invitons les lecteurs à se reporter à notre publication dans la littérature professionnelle (Andro, 2021). L'expérimentation décrite et analysée ici a eu lieu dans un cadre professionnel, au sein des Services du Premier ministre, entre avril 2020 et octobre 2022.
Le CMS libre WordPress a été choisi pour cette expérimentation car c'est le CMS le plus répandu dans le monde, très loin devant Joomla et Drupal. WordPress représente pas moins de 42,5 % des sites web (W3Techs, 2021). Il offre la plus importante communauté de développeurs et, par conséquent, une grande diversité des plugins, notamment autour des technologies RSS. Le dépôt de nom de domaine et l'hébergement en https ont été effectués auprès de OVH pour près de 40 euros par an.
Au coeur de la plateforme, le plugin Feedzy permet de publier automatiquement sous forme d'articles WordPress les flux RSS de n'importes quelles sources et les articles sélectionnés par les veilleurs depuis Inoreader, KB Crawl, Sindup et n'importe quel autre outil de veille. Les contenus envoyés par les flux RSS de sortie de ces outils sont ainsi récupérés sous WordPress automatiquement au statut publié ou brouillon, avec le nom du veilleur et sa catégorie. De manière complémentaire, la plateforme dispose également de fonctionnalités de veille collaborative et permet également de publier des contenus à l'unité avec son login ou via un formulaire web (plugin "User Submitted Posts", grâce à un bouton (plugin "Press this"), grâce à des envois par courriels (plugin "Postie").
Une grande diversité de thèmes et de feuilles de styles compatibles avec la lecture sur smartphones sont disponibles via WordPress ainsi que de multiples fonctionnalités : des fonctionnalités sémantiques via "Darknetcorp Auto Tag Category", iconograpiques via "Quick

Featured Images”, de diffusion via “Newsletter” ou “Subscribe to Category”, de statistiques via “Matomo Analytics”, de veilles événementielles via “The Events Calendar”, de forums de discussion via “bbPress”, de sécurité via “Restricted Site Access” ou de sauvegardes via “UpdraftPlus”.

Concernant les contenus publiés sur la plateforme, seuls sont publiés des synthèses ou des résumés d’articles ou encore de courts extraits de textes entre guillemets et en mentionnant systématiquement les auteurs et les sources. Les images publiées sur la plateforme sont des images libres de droits trouvées via https://search.creativecommons.org ou via la filtre Google image « Licence Creative Commons ». Les contenus originaux publiés sur cette plateforme sont diffusés sous licence CC BY. Les contenus crawlés via KB Crawl ou Sindup ont fait l’objet d’une redevance payée au Centre Français d’exploitation du droit de Copie. Ces contenus, utilisés pour la seule surveillance, ne sont jamais republiés. Pour maintenir un bon niveau de qualité de la plateforme, tous les contenus qui sont publiés sur la plateforme ont été intégralement lus, sélectionnés, catégorisés, étiquetés et, de préférence, analysés par des veilleurs.

Figure 1 : capture d’écran de la plateforme

Analyses

Une veille collaborative en mode pull

Dans les services, les veilles sont traditionnellement diffusées via des newsletters envoyées par courriels en mode “push”. C’est le producteur de contenus de veille qui les diffuse vers des consommateurs de manière irréversible, périodique et verticale, à l’instar de la culture imprimée. Il lui est également difficile de mesurer et de connaître la consommation qui est faite des veilles. Pour les veilleurs, il est difficile de connaître les veilles produites par d’autres services et de développer des collaborations et des mutualisations. Des silos informationnels se forment, des rivalités et des concurrences peuvent se développer au détriment des intérêts des utilisateurs. Le veilleur risque de se couper de ses clients et de s’enfermer (Texier, 2017).

Avec la plateforme de veille, un mode de diffusion “pull” a été expérimenté de manière complémentaire au mode de diffusion traditionnelle. Ainsi, c’est le consommateur de contenus de veille qui va vers l’information de manière volontaire, libre, immédiate, horizontale et interactive puisqu’il peut interagir avec elle et surtout être lui-même producteur de contenus de veille. Ainsi, la veille collaborative permet surtout à n’importe quel utilisateur de la plateforme de partager très librement sur cette plateforme de l'information avec ses collègues de 5 manières différentes :

1- Sans avoir de login, via un simple formulaire web.

2- A partir de son login et de son mot de passe.

3- En l'envoyant automatiquement depuis un agrégateur de flux RSS comme InoReader (solution utilisée principalement par les veilleurs).

4- En republiant un article directement depuis une adresse web.

5- En envoyant l'article à une adresse email spéciale.

A la différence de la culture imprimée, tous les contenus peuvent être modifiés à tout moment. Et à la différence des newsletters, les contenus de veille peuvent être facilement consultés sur smartphone dans les transports en communs. La consommation qui est faite des veilles peut être connue de manière détaillée afin de pouvoir mieux adapter l’offre à la demande. La plateforme permet de mieux capitaliser les contenus de veille, de les rendre recherchables en texte intégral et aussi de pouvoir produire des analyses textuelles de ce corpus de veille. Enfin, la plateforme permet de connaître les veilles produites par d’autres services et de mieux les partager au-delà des silos informationnels. Des collaborations et des mutualisations deviennent possibles. Un meilleur partage du travail devient réalisable. Les économies d’échelle permettent aux veilleurs de passer de la transmission d’informations brutes à leur analyse.

La veille collaborative permet de partager l’information “de chacun selon ses moyens, à chacun selon ses besoins” (Blanc, 1839). Pour prendre une analogie dans le domaine du commerce, si les commerçants choisissent de se réunir sous les mêmes halles d’un marché le même jour, c’est pour partager des frais, pour mieux répondre aux besoins de leurs clients et pour bénéficier d’externalités positives. En effet, le boulanger vendra peut être du pain au client du fleuriste et inversement. Il en est de même pour les veilleurs qui partagent une plateforme de diffusion de leurs veilles.

Du point de vue du consommateur (lecteur)		
	Newsletters envoyées par courriel	**Plateforme web**
Lecture nomade	Difficile (messagerie professionnelle)	Adaptée au smartphones, notamment dans les transports
Mode de diffusion	Push : du producteur vers le consommateur dont la messagerie peut être encombrée	Pull : du consommateur libre et volontaire vers le producteur
Périodicité de publication	Quotidienne, hebdomadaire, mensuelle…	Immédiate
Ouverture sur d'autres veilles	Difficile puisqu'il faudrait s'abonner à une multitude de newsletters	Possible grâce à la structure hiérarchique de la plateforme qui permet de voir toutes les veilles qui intéressent les services en page d'accueil ou de ne lire que des catégories ou sous catégories plus spécifiques

Du point de vue du producteur (veilleur)		
	Newsletters envoyées par courriel	**Plateforme web**
Collaboration entre producteurs	Perfectible : silos informationnels, veilles en doublons	Possible : répartition du travail de veille, partage d'articles et d'une bibliothèque d'images libres de droit
Collaboration avec le consommateur	Verticalité : le consommateur ne peut commenter les contenus ni produire des articles	Horizontalité : le consommateur peut commenter les contenus et en produire lui-même (contenus originaux et renseignement humain par exemples)
Réversibilité	Aucune : les contenus ne peuvent être modifiés après	Totale : à tout moment, les contenus peuvent être

	envoi de la newsletter (culture imprimée)	modifiés (culture web)
Connaissance de la consommation	difficile	détaillée via les statistiques de consultation
Capitalisation des contenus	difficile	adaptée pour rechercher au sein des contenus capitalisés
Analyse des contenus	difficile	Analyses intellectuelles rendues possibles grâce aux économies d'échelle (meilleure répartition du travail, contenus générés par les utilisateurs, automatisation de la diffusion…) et analyses textuelles (text mining) grâce à la capitalisation des contenus.

Du point de vue du commanditaire (direction, actionnaire, contribuable…)		
	Newsletters envoyées par courriel	**Plateforme web**
Coût, efficience	perfectible (silos informationnels)	mutualisation permettant d'améliorer le rendement des veilles (analyses plus poussées par exemple)

D'après une enquête auprès d'experts internationaux de la veille (Kriaa-Mdhaffer, 2018), il ressort que "les synergies et collaborations entre les membres des cellules de veille sont capitales pour l'échange d'informations pour résoudre des problèmes" (pour 94,3% des répondants). Il s'avère aussi que "la synergie entre les cellules de veille peut être renforcée par l'usage des TIC" pour 97,2 % d'entre eux et que "plus les équipes de veille utilisent les TIC, meilleure est la collaboration pour 81,4 % d'entre les répondants.

Vers un réseau d'organisations apprenantes

Nos organisations élèvent leur niveau de conscience via la formation, la communication et la culture managériale et via la pleine conscience de leur environnement dont elles doivent analyser la complexité croissante et anticiper le changement rapide (Senge, 2015). Il devient vital que l'intelligence individuelle devienne plus collective. L'information doit pouvoir circuler de manière fluide dans l'organisation comme dans un système nerveux pour déclencher une réaction rapide et vitale ou pour être capitalisée sous forme de connaissances susceptibles

ensuite d'alimenter une vision et une stratégie. La collaboration et le partage d'une vision commune est indispensable à l'efficacité. Les modèles de management évoluent vers plus de collaboration et d'agilité. La fonction veille doit accompagner ce changement et combiner intelligence collective, artificielle et économique.
On peut identifier divers niveaux de maturité vers l'organisation apprenante. Les veilles personnelles, non automatisées sur des sources ouvertes qui existent dans toutes les organisations peuvent laisser progressivement la place à des outils gratuits puis à une plateforme professionnelle portée par une cellule de veille (Chalus-Sauvannet, M.-C., 2021). L'intelligence peut devenir plus distribuée et collaborative, chaque collaborateur devient acteur de la veille. Les contenus peuvent être diffusés et intégrés aux systèmes d'information où interagissent déjà les collaborateurs de l'organisation (France Processus, 2022). Cette évolution doit être accompagnée par des sponsors et par la direction de l'organisation.

Un projet développé de manière agile sous une forme originale de collaboration public / privé

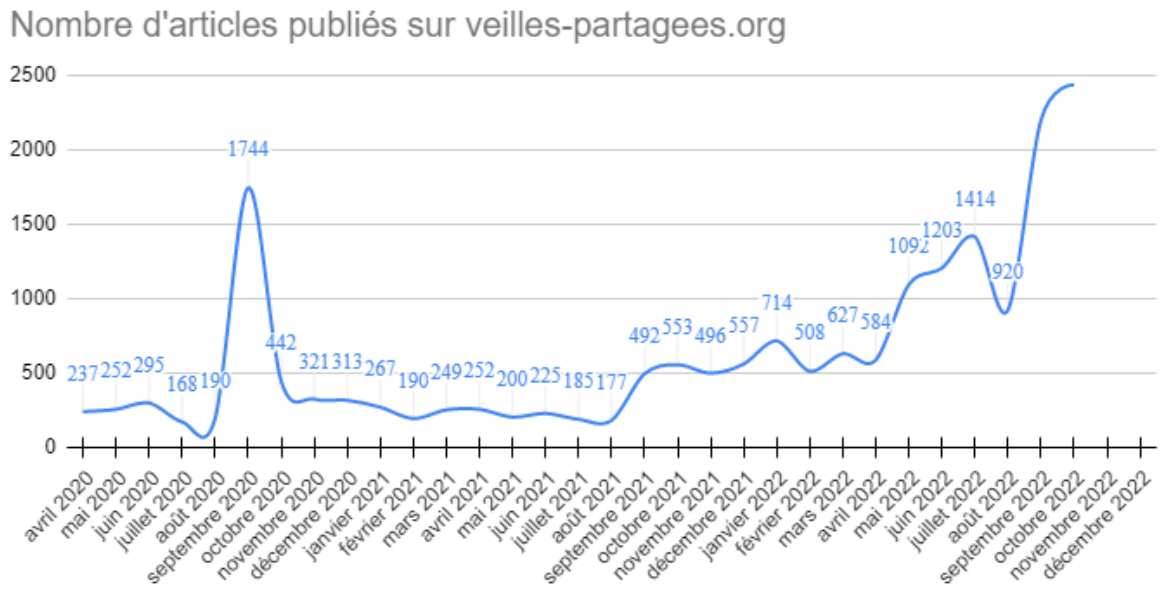
La plateforme a été développée selon un modèle de "user innovation" (Von Hippel, 2005). Ce n'est pas le producteur qui a construit une offre pour le consommateur mais le consommateur qui a bricolé un produit répondant à ses besoins et ce produit a ensuite été repris par un éditeur.
Traditionnellement, ce type de projet d'acquisition d'une plateforme de veille, aurait pu débuter par une phase d'audit et d'analyse des besoins, il aurait ensuite été suivi par la rédaction d'un cahier des charges sans relation directe et immédiate avec la réalité concrète. Le cahier des charges aurait ensuite été validé par un comité de pilotage, par la publication d'un appel d'offres, son dépouillement et le développement spécifique et/ou l'achat d'une plateforme de veille pour quelques dizaines de milliers d'euros après quelques années de conduite du projet avec un possible échec à la clé, l'outil ne correspondant pas bien aux capacités numériques des veilleurs et/ou aux besoins de leurs usagers.
Dans le cas présent, l'expérimentation est partie de l'intuition qu'il était possible de développer une plateforme de veille avec WordPress suivie de sa réalisation concrète en quelques jours et pour près de 40 euros par an sous la forme d'un premier prototype expérimental. La plateforme a ensuite été progressivement adoptée par un nombre croissant de veilleurs et adaptée à leurs besoins et à ceux des usagers grâce à des plugins. Après publication d'une description détaillée de l'expérimentation dans la littérature professionnelle (Andro, 2021), l'expérimentation a été reproduite par la société Sindup qui cherchait, de son côté, à proposer ce type de plateforme WordPress à ses clients. La plateforme a ainsi pu être reprise, chartée et sécurisée par cet éditeur de technologies de veille.
Les ressources financières ainsi épargnées ont ainsi pu être concentrées sur les technologies de surveillance notamment de web crawling et API qu'il reste aujourd'hui préférable d'externaliser.

Analyses des usages

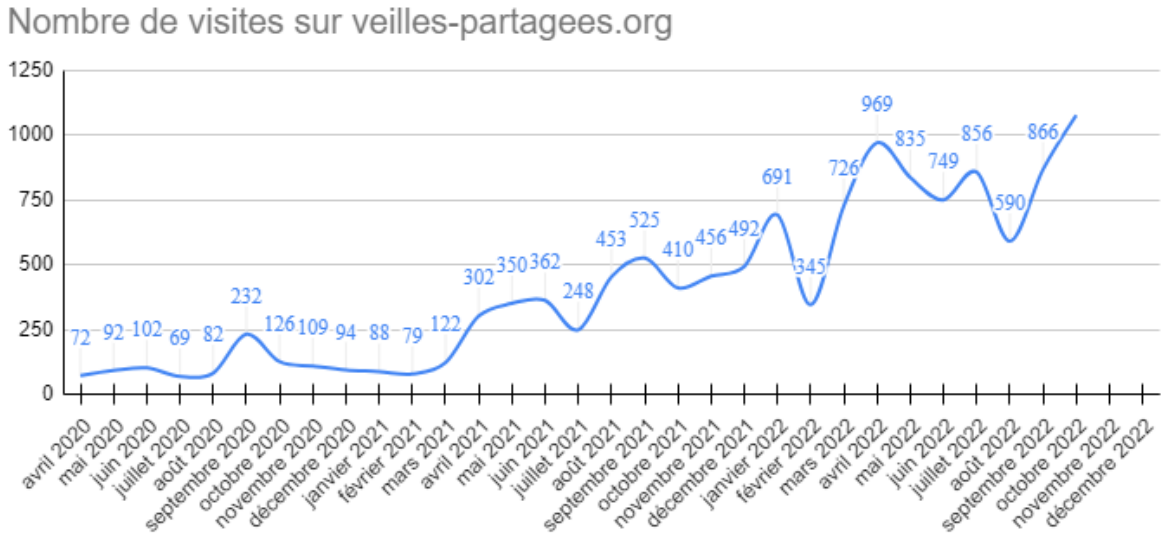
Entre avril 2020 et avril 2021, le prototype est resté accessible en open access sans restrictions, diffusant essentiellement une veille sur la veille à partir de sources ouvertes. Le besoin d'utiliser la plateforme WordPress afin d'élargir à d'autres veilles au sein du réseau nous a contraint à restreindre l'accès au site dès avril 2021. Le besoin de sécuriser, de mettre en forme et d'officialiser son existence nous a conduit à construire un partenariat avec l'éditeur Sindup en septembre 2021.

Sur la première période "Open access" (avril 2020 - avril 2021), 937 visiteurs ont généré 3907 sessions d'une durée moyenne de 8 minutes 54 secondes et ont vu 20 395 pages web. Les données de ce trafic web ont été obtenues grâce à Google Analytics. A partir de la restriction d'accès par adresses IP et logins, Google Analytics ne pouvait plus avoir accès à nos données. Nous avons donc eu recours au logiciel Matomo Analytics afin de connaître les usages de la plateforme. En cela, nous avons involontairement anticipé une mise en demeure de la Commission nationale de l'informatique et des libertés qui considère qu'"*il existe donc un risque pour les personnes utilisatrices du site français ayant recours à cet outil et dont les données sont exportées."* et qui "*constate que les données des internautes sont ainsi transférées vers les États-Unis en violation des articles 44 et suivants du RGPD*" (Cnil, 2022). Quoi qu'il en soit, sur la deuxième période en "Accès restreint" (mai 2021 - février 2022), 4 462 visites ont consulté 14 942 pages web d'après les données extraites via Matomo.

Statistiques de production



Statistiques de consommation



Les Services du Premier ministre comptaient 3 238 agents publics (dont 2018 cadres A et A+) au 31 décembre 2020 (Bilan social des Services du Premier ministre, 2020). Avec 1075 visiteurs uniques sur le mois d'octobre 2022, c'est un peu plus de 33 % de l'effectif total qui a fréquenté la plateforme sur cette période d'un mois. Si on considère que la veille est principalement destinée aux cadres A et A+, on peut estimer en avoir touché plus de la moitié sur la période.

La restriction d'accès ne s'observe pas sur le diagramme précédent qui montre la dynamique temporelle des articles publiés, le nombre de visites et celui des pages vues. En effet, le caractère expérimental et confidentiel de la plateforme sur la première période n'avait pas permis à la plateforme de générer un très fort trafic web. Par ailleurs, le réseau de veille s'est progressivement élargi, passant de 58 membres en avril 2020 à 143 en février

2022. Enfin, l'utilisation de la plateforme pour y diffuser un nombre croissant de veilles a probablement contribué à atténuer la chute de trafic web lié à la restriction d'accès. La situation de crise de la covid-19 et son alternance de confinements, de télétravail, de travail hybride… a également été de nature à perturber nos analyses d'usages.
Ainsi, concernant les analyses d'usages, nous considérons que la question de recherche la plus intéressante en termes de production de connaissances consiste à vérifier l'hypothèse selon laquelle des contenus de qualité génèrent davantage de trafic web que des contenus bruts simplement rediffusés.
Afin de pouvoir comparer le trafic web générés par des contenus mis en ligne à des dates différentes et au sein de collections de volumes différents, nous avons fait le choix de retenir un indicateur de nombre de visites par an et par article et de comparer ces indicateurs pour les 3 corpus suivants correspondant à 3 niveaux de qualité :

1. **Veilles automatiques** : contenus de veilles complètement automatisés comportant le titre de l'article original, le lien URL pour y accéder et une image automatique par catégorie de veille
2. **Simple curation humaine** : Contenus de veilles issus d'une sélection non automatique et comportant aussi le titre, l'URL et une image libre de droit d'illustration
3. **Synthèses et contenus originaux** : Articles ayant été lus, résumés ou synthétisés et/ou dont un extrait représentatif a été choisi sous la forme d'une citation, mais aussi contenus originaux (comptes-rendus de réunions, retour d'expériences, expérimentations, renseignement humain…). En plus de ces contenus, on trouve naturellement aussi le titre, l'URL éventuelle et une image libre de droit.

Nos articles n'étant pas catégorisés en fonction de ces 3 niveaux, nous avons fait le choix de ne pas analyser les 9 873 articles publiés sur la plateforme mais de nous contenter d'échantillons sur l'année 2021 majoritaires au sein de certaines catégories thématiques de la plateforme.

	Nombre d'articles en 2021	**Nombre cumulé de vues en 2021**	**Nombre moyen de vues par articles en 2021**
Veilles automatiques	1956	229	0,12
Simple curation humaine	477	119	0,25
Synthèses et citations	592	862	1,46

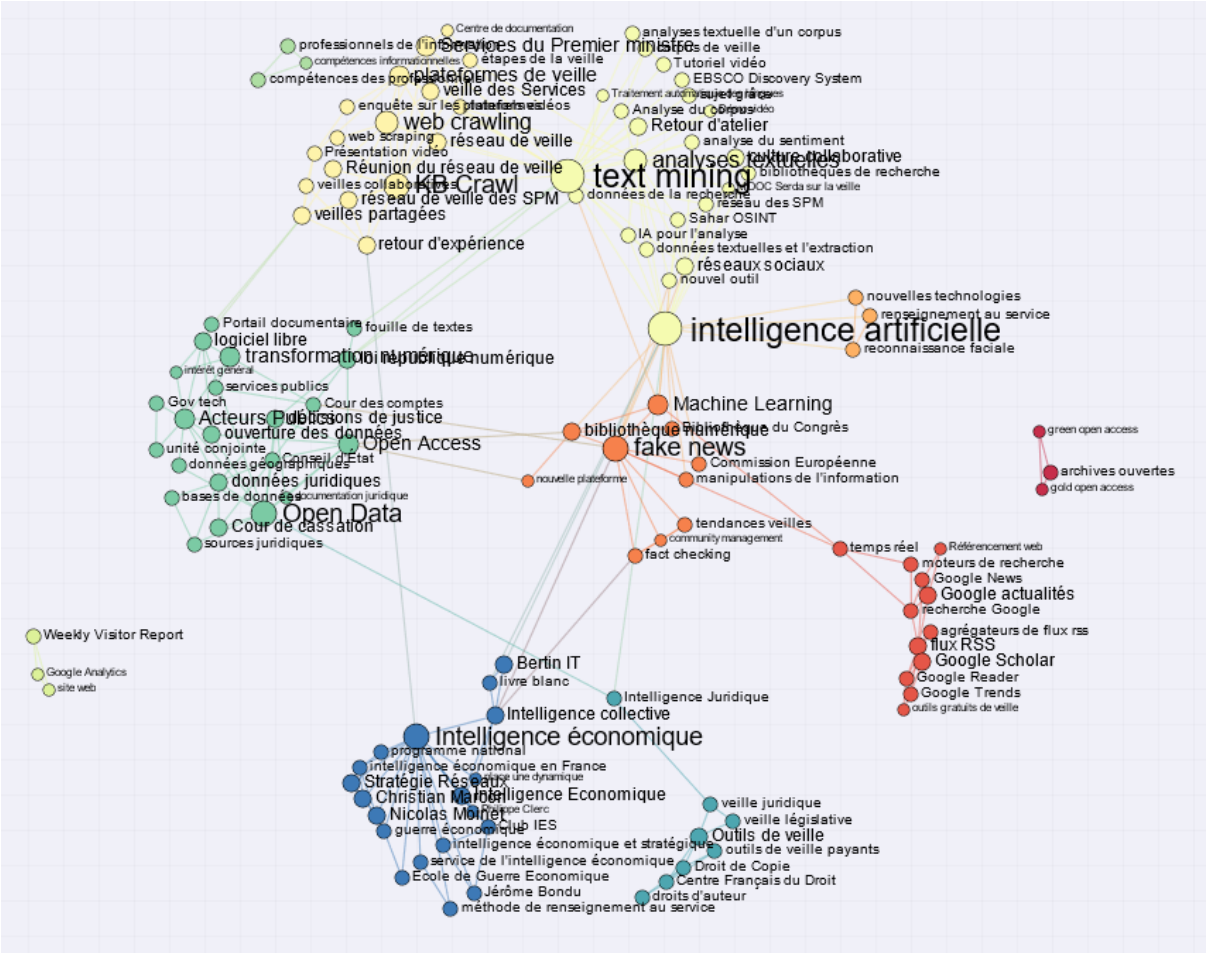
Sur un corpus restreint, nous vérifions ainsi l'hypothèse qui suppose que la veille à plus haute valeur ajoutée génère aussi un trafic web plus important et répond aussi mieux aux besoins des usagers. Dit autrement, plus le contenu publié sur la plateforme de veille a nécessité un travail intellectuel et plus il semble générer de trafic web.
En effet, le métier de veilleur ne consiste plus à noyer les commanditaires d'une masse de contenus non lus comme des robots peuvent désormais le faire. Le métier de veilleur consiste aujourd'hui à filtrer les contenus détectés pour ne communiquer que les plus

pertinents d'entre eux. En cela, le veilleur fait gagner un temps précieux à ses collègues. Mais ce travail lui-même est susceptible d'être automatisé grâce à de l'intelligence artificielle apprenant à identifier parmi les nouveaux documents détectés s'ils ressemblent davantage à l'"empreinte sémantique" des documents du corpus de documents validés précédemment par le curateur humain ou plutôt à des documents du corpus des rejetés.
Le travail du veilleur tend donc à utiliser au maximum les possibilités d'automatisation et d'intelligence artificielle qui sont à sa disposition pour se concentrer sur un travail de lecture, de résumé, d'analyse de contenus et, pourquoi pas, de production de contenus originaux et d'expertise.

Analyses textuelles des contenus avec CorTexT

Le travail d'analyse ne peut pas encore être automatisé grâce à de l'intelligence artificielle. Il devient néanmoins, dès à présent, possible d'utiliser les technologies de text mining afin de faciliter ce travail. Face à l'infobésité, l'intelligence artificielle CorTexT, développée par le recherche scientifique française[2], nous a ainsi permis d'identifier automatiquement les termes conceptuels et les entités nommées de lieux géographiques, de noms d'organisations et de noms de personnes au sein des corpus textuels de nos veilles publiées sur la plateforme. Il nous a ensuite été possible d'en donner une représentation cartographique avec une visualisation objective et "data driven" des occurrences de ces entités et de leurs cooccurrences au sein des articles de veille.
Depuis l'ouverture de cette plateforme il y a 2 ans, nous avons publié près de 1000 articles pour la seule veille sur la veille. Afin d'analyser objectivement ce corpus de veille, nous avons donc, grâce à CorTexT, identifié les concepts, puis nous les avons cartographié par occurrences (taille des noeuds de concepts) et cooccurrences (liens entre les noeuds concepts regroupés en clusters).
Voici la cartographie obtenue à partir des concepts identifiés au sein du corpus de veille :

[2] Institut national de recherche pour l'agriculture, l'alimentation et l'environnement et Institut Francilien Recherche Innovation Société



Globalement, on retrouve bien dans cette cartographie les différentes catégories de notre veille métier de veilleur :

- Technologies
  - Outils de veille
    - Technologies RSS et Inoreader
    - Web crawling et web scraping
  - Datavisualisation, text mining et intelligence artificielle
- Transformation numérique, intelligence économique et management
  - Open access et open data
  - Veille juridique et cadre juridique de la veille
  - Cybersécurité et cyberdéfense
- Editeurs et ressources documentaires
- Evénements veilleurs

Discussion

Sans revenir sur les discussions entre partisans du logiciel libre et du logiciel dit propriétaire qui, dans le domaine de la veille, tourne souvent autour de la technologie RSS (Barthes, 2020), notre approche se veut pragmatique, faisant le choix du logiciel libre là où il nous

semble plus performant (CMS WordPress) et le choix du logiciel propriétaire dans d'autres cas (web crawling avec KB Crawl, veille avec Sindup…)
Concernant la veille collaborative, la principale difficulté à son développement dans les organisations provient de cultures de management fondé sur la rareté de l'information. "Ceux qui savent ne partagent pas pour conserver leur statut" (Bondu, 2017), de structures pyramidales et cloisonnées ou encore de freins cognitifs et de déficit de culture collaborative. Le développement des entreprises opales, dont le concept renvoie au dernier stade de l'évolution organisationnelle et humaine (Frédéric Laloux, 2014), provoque un changement de paradigme qui rompt progressivement les freins habituellement rencontrés dans les organisations plus traditionnelles.
Cette expérimentation mériterait d'être reproduite dans d'autres organisations afin de conforter les résultats obtenus. Une vidéo de démonstration de la mise en place d'une telle plateforme depuis l'achat du nom de domaine jusqu'à l'envoi de contenus vers WordPress depuis un outil de veille sera proposée prochainement afin d'en permettre plus facilement la reproduction. Ce type de plateforme développée avec du logiciel libre pourrait ainsi faciliter le recours à la veille collaborative dans les organisations.

Conclusion

Le métier de veilleur a connu des gains en efficience considérables grâce à l'automatisation quasi complète de la surveillance des sources avec les API, le web crawling et la technologie RSS. Aujourd'hui, la diffusion semi automatique de contenus et leur analyse assistée par de l'intelligence artificielle permettent de supprimer les tâches les moins épanouissantes en termes de développement humain.
En matière de développement de plateforme de diffusion de la veille, une séparation trop forte entre théorie du client-décideur-métier et pratique du prestataire-développeur nous semble susceptible de générer un effet tunnel, le client demandant des projets irréalistes et sources de quiproquos et de déceptions. S'inspirant des méthodes agiles et du concept de "user innovation", le choix a été fait d'une approche plus pragmatique, collaborative, ouverte et centrée sur les besoins de l'usager.
Une relation plus partenariale entre veilleurs et éditeurs de technologies de veille a également été expérimentée à cette occasion. Nous espérons que cette expérimentation sera reproduite par d'autres veilleurs et d'autres éditeurs et qu'une communauté de professionnels utilisant WordPress afin de diffuser leurs veilles pourra être construite sur cette base afin de mieux partager les bénéfices d'expérimentations diversifiées.

Références

Andro, M., Bigoy, C., Terrière, M., Thébault, D. (2021). Comment développer gratuitement une plateforme collaborative de veille avec le logiciel libre WordPress ?. ADBS, 2021, 16 p. (https://www.adbs.fr/groupes/adbs-info/comment-developper-288932)

Barthes, Emmanuel (2020). Les avantages des outils de veille payants: Un vieux débat. https://www.precisement.org/blog/Les-avantages-des-outils-de-veille-payants.html

Bilan Social 2020 des Services du Premier ministre. https://www.documentation-administrative.gouv.fr/adm-01859571/document

Blanc, Louis (1839). Organisation du travail. https://gallica.bnf.fr/ark:/12148/bpt6k24230t

Bondu, Jérôme (2017). Une veille collaborative… ou rien. https://www.cairn.info/revue-i2d-information-donnees-et-documents-2017-3-page-12.htm

Breucker P., Cointet J., Hannud Abdo A., Orsal G., de Quatrebarbes C., Duong T., Martinez C., Ospina Delgado J.P., Medina Zuluaga L.D., Gómez Peña D.F., Sánchez Castaño T.A., Marques da Costa J., Laglil H., Villard L., Barbier M. (2016). CorTexT Manager (version v2). URL: https://docs.cortext.net

Chalus-Sauvannet, M.-C. (2021). Veille stratégique à l'usage des décideurs: Anticiper les opportunités pour conduire des stratégies innovantes. EMS

Commissariat Général du Plan - Intelligence économique et stratégie des entreprises- Travaux du groupe présidé par Henri Martre (1994). https://www.entreprises.gouv.fr/files/files/directions_services/information-strategique-sisse/rapport-martre.pdf

Jérôme Bondu, Cahiers de la Documentation - Bladen voor Documentatie (2020), Facteurs clés de succès d'une veille collaborative https://www.abd-bvd.be/wp-content/uploads/2020-1_Bondu.pdf

Commission nationale de l'informatique et des libertés (2022). Utilisation de Google Analytics et transferts de données vers les États-Unis : la CNIL met en demeure un gestionnaire de site web. https://www.cnil.fr/fr/utilisation-de-google-analytics-et-transferts-de-donnees-vers-les-etats-unis-la-cnil-met-en-demeure

France Processus (2022). Processus et transversalité(s) - Vers un nouveau management. Afnor

Jéromine Gay (2021). Travail de Bachelor en vue de l'obtention du Bachelor HES, Proposition d'un dispositif de veille stratégique au sein d'un organisme de soutien à l'économie régionale valaisanne. Quel mode d'organisation de la veille et quel outil choisir pour une structure d'une dizaine de personnes ?
https://sonar.ch/documents/315071/files/GAY_JEROMINE_ID_TB2021.pdf

Stéphane Goria, ISTE Editions (2017), Méthodes et outils de veille créative

Kriaa-Mdhaffer, Salima & Janissek-Muniz, Raquel & Lesca, Humbert. (2018). TIC et synergie dans les cellules de Veille Stratégique.
https://www.lume.ufrgs.br/bitstream/handle/10183/187354/001080012.pdf

Frédéric Laloux (2014). Reinventing Organizations - La version résumée et illustrée du livre phénomène qui invite à repenser le management. Diateino

Moreau, I. & Rodrigue, J. (2008). La situation de la veille informationnelle dans les organisations gouvernementale. Documentation et bibliothèques, 54(4), 273–283.
https://doi.org/10.7202/1029190ar

Senge, P. Arnaud, B., Gauthier A. (2015). La cinquième discipline - Levier des organisations apprenantes, Eyrolles, 2015.

Texier, Bruno (2017). Veilleurs, sortez de votre bulle informationnelle !
https://www.archimag.com/veille-documentation/2017/03/23/veilleurs-sortez-bulle-informationnelle

Von Hippel E., Democratizing innovation, MIT Press, Cambridge, 2005.

W3Techs (2021) Usage statistics and market share of WordPress.
https://w3techs.com/technologies/details/cm-wordpress